RATIONAL INATTENTION EXERCISE

The hard part of this problem set was the last. If you got numerical results that were correct, or nearly correct, you got a "4" grade. If your numerical methods remained stable into the region where the solution shifts from having three points of support to having two points of support, or you recognized that your methods broke down there (and that this point is not the point where the solution has a single point of support), you got a 5. I thought this problem was very challenging, and the performance of the class was remarkably good.

Suppose a monopolist faces the demand curve $y = \pi^{-\theta}$, where y is quantity sold and π is the price. He faces a randomly varying cost c, which takes on the three values $\{1, 2, 8\}$ with probabilities $\{.4, .4, .2\}$, respectively. He wants to maximize expected profits $y \cdot (\pi - c)$. He faces a cost of mutual information between c and his choice of π, i.e. $H(c, \pi)$, of α per bit.

(1) Show that, whatever the joint distribution of π and c, at any value of π that has positive probability,

$$\pi = \sum_j E\left[\frac{\theta}{\theta - 1} c_j \mid \pi\right] .$$

The problem is

$$\max_{p_{ij}, \pi_i} \sum_{i,j} p_{ij} \pi_i^{-\theta} \cdot (\pi_i - c_j)$$
$$- \alpha \cdot \left(\sum_{i,j} p_{ij} \log(p_{ij}) - \sum_{i,j} p_{ij} \log(\sum_k p_{ik}) \right)$$
$$\text{subject to} \qquad \sum_i p_{ij} = q_j \, , \text{ all } j \, ,$$

where q_j is the probability of cost c_j.

The FOC's with respect to $\{\pi_i\}$ give us

$$\sum_j \left((1 - \theta) \pi_i^{-\theta} - \theta \pi^{-\theta - 1} c_j \right) p_{ij} = 0 \, .$$

Solving this for π produces

$$\pi_i = \frac{\theta}{\theta - 1} \frac{\sum_j p_{ij} c_j}{\sum_j p_{ij}} \, ,$$

which is what you were asked to show.

(2) Suppose the monopolist collects no information about c. What is the joint distribution of c and π in that case?

In this case the monopolist chooses a single value for π, which will be

$$\pi = \sum_j q_j \frac{\theta}{\theta - 1} c_j = \frac{\theta}{\theta - 1} E[c_j] \,.$$

The joint distribution makes π and c independent, with π non-random, in other words.

That there is only one value of π chosen follows from the expression for optimal π above, together with the fact that, with mutual information between π and c zero, the two must be independent, and therefore expected c conditional on π must be the same for every π.

That mutual information is zero when π and c are independent is easy to see from the definition of mutual information. That zero mutual information implies independence requires a little more argument. If we fix the marginals, so $\sum_i p_{ij}$ and $\sum_j p_{ij}$ are both fixed, and take FOC's, we see that minimal mutual information implies $\log p_{ij} = -1 + \lambda_i + \mu_j$ for all i, j, where λ_i and μ_j are the Lagrange multipliers on the constraints fixing the marginals. Then it is easy to see this implies that each p_{ij} is the product of the marginal probabilities, i.e. that we have independence.

(3) Suppose the monopolist's cost of information is so low that he observes c perfectly. What is the joint distribution in that case? What is the mutual information in that case?

In this case π and c are perfectly correlated, so $p_{ii} = 1$, $i = 1:3$ and $p_{ij} = 0$, all $i \neq j$. The value of π_j is just that in the formula for the optimal π, above, without the need for an expectation sign. The mutual information is the entropy of the c_j distribution, i.e. 1.52 bits.

(4) Suppose that π can take on at most three distinct values, $\{\pi_i, i = 1, \dots, 3\}$. (It probably can be proved that this is implied by optimal behavior.) Find a solution that involves some information collection, but not perfect observation. (This can be done mostly with matrix algebra for a solution where every value of π has positive probability, by using the first order conditions.)

(5) Find values of α low enough to make the full-information solution optimal and high enough to make the no-information solution optimal.

I meant to make this a purely numerical question, but forgot to give you a value for θ. Of course we need $\theta > 1$, as otherwise the solution is always to set $\pi = \infty$. Let's take $\theta = 2$. Furthermore, solving this requires much more than "matrix algebra", even when all p_{ij}'s are positive. The FOC's in that case do give two equations that, when solved sequentially, are just a pair of linear systems in κ and the marginal probabilities r_i of the π_i's. But the coefficients in these equations are functions of the π_i's, which in turn depend on the r_i's. Trying to solve directly with a nonlinear solver is possible, but if the solution does not have full support (three non-zero probabilities for distinct values of π) or information cost is very low, the solver will misbehave. There is an algorithm, called the Blahut-Arimoto algorithm, that computes solutions to problems of this form by iterating to a fixed point. Code that does it is posted with this answer sheet and described below.

Using it, I've found that an α of .25 in nat units is enough to make the solution reduce to the no-information case, i.e. a single price of $2E[c_j] = 5.6$ no matter what c_j is.

An α of .04 is low enough to give a solution with nontrivial weight on three distinct prices: approximate prices of 2.1, 3.2 and 14.1 with probabilities .19, .58, .23. In this case the conditional distributions for cost given the three prices are the rows of the matrix below:

$$\begin{bmatrix} 0.93 & 0.07 & 0.00 \\ 0.39 & 0.61 & 0.00 \\ 0.01 & 0.15 & 0.84 \end{bmatrix} .$$

This corresponds to mutual info of .74 bits.

An α of .1 leads to a solution with two points of support: $\pi \in (2.9, 9.9)$ with probabilities .61 and .39 and with conditional distributions the rows of

$$\begin{bmatrix} 0.56 & 0.44 & 0.00 \\ 0.14 & 0.34 & 0.52 \end{bmatrix} .$$

This corresponds to mutual information of .36. bits.

Finally, if α is as low as .01, one gets essentially the full information solution, with prices 2, 4, and 16 and the joint distribution concentrated on the diagonal. While there is an α value so high that no information at all is collected, there is no value so low that one gets *exactly* the full information solution. My numerical solutions show conditional probabilities of .998 and higher on the diagonal, but not ones.

The FOC's w.r.t. p_{ij} imply that, at values of π_i with positive probability,

$$\pi_i^{-\theta}(\pi_i - c_j) = \lambda_j - \alpha \left(\log p_{ij} + 1 - \log\Big(\sum_j p_{ij} \Big) \right) .$$

Exponentiating this gives us

$$p(j \mid i) = \kappa_j \exp(\alpha^{-1} \pi_i^{-\theta}(\pi_i - c_j)) . \qquad (*)$$

Let $M(\alpha) = [\exp(\alpha^{-1}\pi_i^{-\theta}(\pi_i - c_j))]$ (i.e. the matrix with i, jth element as given in that expression). Then, because conditional probabilities sum to one, we have

$$M(\alpha, \pi) \cdot \kappa = \mathbf{1} , \qquad (\dagger)$$

where "$\cdot$" is matrix multiplication. If we solve this for κ, then we have

$$r \cdot M(\alpha, \pi) \cdot \mathrm{diag}(\kappa) = q , \qquad (**)$$

where q is the marginal probability for the cost. This can be solved for r. But of course the equation for $p(j \mid i)$, $(*)$ above, makes π in turn a function of κ and π that is nonlinear in π. The Blahut algorithm iterates on these calculations, starting with a guess for r and π, solving for r as above, using the new r to calculate the corresponding new π, etc. This turns out to converge nicely. When one uses more points of support for the distribution of π than necessary, the converged result will show some identical values of π, corresponding to identical $p(i \mid j)$ distrbutions, with the marginal probabilities that have to be summed.

Hint: A general solution will require some numerical optimization. If you take that route, you will need to do constrained optimization, e.g. by introducing a penalty function for $p(c, \pi)$ values extremely close to zero.